心理科学进展 2018, Vol. 26, No. 12, 2113–2128
Advances in Psychological Science DOI: 10.3724/SP.J.1042.2018.02113

·元分析(Meta-Analysis)·

校园反欺凌项目干预效果及影响因素：元分析和 GRADE 证据质量*

赵陵波[1] 赖丽足[2] 林羽中[1] 赵春晓[2] 任志洪[2]

([1]福州大学人文社会科学学院，福州 350108) ([2]青少年网络心理与行为教育部重点实验室，华中师范大学心理学院，湖北省人的发展与心理健康重点实验室，武汉 430079)

摘 要 通过文献检索和筛选，选取了 43 个校园欺凌的干预研究，使用 CMA 3.0 分析干预效果及影响因素，并使用 GRADE 系统对所有结局指标的证据质量进行评估。结果显示，针对欺凌者的干预项目具有中等效果量($g = 0.57$, $p < 0.05$)；针对欺凌受害者的干预项目使得受害者的受欺凌行为减少($g = 0.42$, $p < 0.05$)，心理健康得到改善($g = 0.40$, $p < 0.01$)；针对所有学生的欺凌干预项目对其欺凌行为的效果 $g = 0.17$ ($p < 0.001$)，对其受欺凌行为效果 $g = 0.19$ ($p < 0.001$)，可能存在出版偏差；学生态度改变的效果 $g = 0.40$ ($p < 0.01$)。除了态度改变的证据质量为中等，其他结局指标的证据都为低或者极低质量。对于所有学生干预项目，干预周期小于 1 学期的效果量大于干预周期大于 1 学期的效果量，教育干预项目效果量小于系统干预效果量，非随机研究效果量大于随机对照研究。直接针对欺凌者或者欺凌受害者的干预效果量中等，但是样本量少，需要进一步的研究证据支持。针对所有学生的欺凌干预项目虽然效果量低，但是仍具有实践意义，其效果量受到干预周期和干预特征的影响。

关键词 校园反欺凌；干预效果；元分析；GRADE
分类号 B844

1 背景

校园欺凌可定义为一个或者多个学生对同伴进行身体、心理的反复迫害或者道德攻击，主要有以下三个特征：欺凌者和受害者力量上的不均等；攻击行为是故意且敌意的；欺凌反复发生(Olweus, 1994a)。欺凌的形式不仅仅包括直接的暴力行为，也包含间接的攻击行为；内容包括身体欺凌(例如踢打等身体暴力行为)，口头欺凌(例如嘲笑，取侮辱性的绰号)，关系欺凌(例如散播谣言，伤害他人名誉)，损坏受害者所有物。并且随着网络技术的发展，网络欺凌成为新的欺凌渠道，其包括电子邮件、短信、视频等(Kowalski, Giumetti, Schroeder, & Lattanner, 2014)。

欺凌形式、渠道的多样性提升了欺凌行为的发生率。一项综合了全球 80 项报告青少年欺凌行为发生率研究的元分析发现：传统的欺凌现象发生率可达到 34.5%，网络欺凌率为 15.5% (Modecki, Minchin, Harbaugh, Guerra, & Runions, 2014)。一项整合了大陆、台湾、香港和澳门两岸三地的调查研究发现，青少年欺凌报告率最高达到 68%，受欺凌报告率最高达到 66% (Chan & Wong, 2015)。欺凌行为会严重影响欺凌者、欺凌受害者的身心健康和发展。研究显示，欺凌者更容易有消极的或者反社会行为(例如逃学，犯罪，物质滥用) (Havik, Bru, & Ertesvåg, 2015; McCuddy & Esbensen, 2017; Ttofi, Farrington, Lösel, Crago, & Theodorakis, 2016)。而欺凌受害者与低自尊、低自我价值感正相关，更容易有焦虑、抑郁、自杀意念等心理问题(Kelly et al., 2015)。因此校园欺凌现象不仅普遍，并且造成严重的危害，鉴于此，

收稿日期：2018-04-13
* 福建省社科基金资助(FJ2017B028)。
通信作者：任志洪, E-mail:ren@mail.ccnu.edu.cn

欺凌的干预和预防逐渐引起政府、教育部门和公众的重视。

自 1983 年, Olweus 在挪威设计了第一个大规模的校园欺凌干预项目起，之后的数个效果评估显示校园欺凌/受害现象下降了 21%~50%左右(Olweus, 1994b, 2005)。随后，越来越多研究者根据不同理论设计了校园欺凌干预项目，并将之应用于中小学学校中。这些干预项目涉及个体层面、学校层面、家庭层面，乃至社区和社会层面(Farrington, Gaffney, Lösel, & Ttofi, 2016)。在个体层面，针对学生对欺凌的知识、态度进行教育；在学校层面，制定一定的校规，或者组建委员会处理欺凌事件；在家庭层面，增加家长对欺凌的认识，增强家长和学校的互动；在社会层面，制定相应的政策应对欺凌事件。

在干预项目之后，大量的实证研究对各种反欺凌项目的干预效果进行了检验，可是得出的结果并不一致，很多综述和元分析都对其进行了梳理。Vreeman 和 Carroll (2007)归纳了 1996 年至 2004 年共 26 个学校的欺凌干预研究，其中有 9 项研究结果显示干预没有显著降低欺凌或者受欺凌行为。Evans, Fraser 和 Cotter (2014)将研究更新到 2013 年，发现测量欺凌行为的 22 个研究中，只有 50%显示干预显著有效，测量受欺凌的 27 个研究中有 67%显示受欺凌行为显著降低。Merrell, Gueldner, Ross 和 Isava (2008)的元分析计算了校园欺凌干预的效果量，指出反欺凌项目效果量很低甚至无效(Merrell et al., 2008)。但也有一些元分析得出校园欺凌干预是有效的(Jiménez-Barbero, Ruiz-Hernández, Llor-Zaragoza, Pérez-García, & Llor-Esteban, 2016; Ttofi & Farrington, 2011)。不难发现，这些元分析结果的不一致可能源于纳入排除标准和效果量计算的差异，有以下几点需要进一步考虑。

第一，纳入研究的实验设计有待考虑。在循证研究中，一般认为随机对照实验(random control trial, RCT)具有较好的信效度。然而对于学校干预而言，要考虑学校的意愿，很难实现随机化的分配。即使随机化，对参与实验的学校数量也有较高的要求。而实际上很多 RCT 研究并非大规模的干预研究，可能并没有达到随机化的效果。较新的一篇元分析，只关注 RCT 研究，纳入了 16 篇研究，虽然一定程度上保证了内部效度，但却忽视了其他的实证证据(Jiménez-Barbero et al., 2016)。而前后测实验组-控制组非 RCT 设计控制了因变量在基线水平上的差异，一定程度上可以减少非随机带来的弊端，这部分的证据也是重要的。尽管也有元分析纳入了非 RCT 的对照研究，但是在计算效果量的时候却只采用了后测数据，忽视了两组的基线水平上的异质性。考虑干预前测数据，计算两组前后测变化量的标准化差异，是更加合理的效果量计算方式(Hoyt & Del Re, 2017)。即便对于 RCT 研究而言，前后测的重复测量也能够有效提高统计效力和检验力(Carlson & Schmidt, 1999)。

第二，对干预项目的分类没有细化。根据干预对象，反欺凌干预可以划分为三类，一类是直接指向欺凌者，旨在减少他们的攻击性或者欺凌行为(Şahin, 2012)；一类是直接指向欺凌受害者，通过心理健康辅导、社交技能训练等提高他们的心理健康水平，降低受欺凌频率(Fox & Boulton, 2003)；还有一类是更为广泛的欺凌干预，针对所有学生，通过教育、纪律、和家长社区的合作等增强学生对欺凌的认识，提升求助的意识，进而学生的态度和行为(Saarento et al., 2015)。以上三类干预中，对于特定对象的干预而言，干预和测量都是直接命中目标对象，需要在干预前进行有效的识别。而对于普遍干预而言，面向的是所有学生，是“预防 + 干预”的结合。因此，在计算效果整体效果量的，三种类型的干预是有必要分开的。

第三，未对证据质量强度作评估或者评价不全面。尽管元分析是综合证据结果的强有力工具，但是并非所有的整体效果量都是有意义的，对于证据质量的强度的评价是必要的。文献质量评价有多种工具，例如 Cochrane 风险偏倚评估、Jadad 量表等，这些评估主要集中于研究设计的质量(Zeng et al., 2015)。GRADE 工作组于 2004 年正式推出一套循证医学证据质量分级和推荐系统，不仅可以评估 RCT 证据，也可用于评估非 RCT 研究证据，已经成为国际接受的标准之一。作为系统评价的一部分，GRADE 不只关注研究的局限性，还关注研究的不一致性、间接证据、结果的精确性、发表偏倚，更加全面的对元分析结果的证据强度进行评估，有利于指导临床实践(Alonso-Coello et al., 2016)。此外，GRADE 系统不仅仅关注单一结果，还关注其他重要的效果衡量指标。因此，本研究纳入了欺凌行为、受欺凌行为、心理健康、态度等因变量作为衡量反欺凌干预项目效果的结

局指标(outcome variables)。

综上，本研究除了更新已有的反欺凌干预研究，还将考虑纳入前后测实验组−控制组干预效果研究，运用前后测数据进行干预效果量的计算。根据测量对象的不同，我们分别计算针对欺凌者的干预，针对欺凌受害者的干预，及针对所有学生的干预效果，并采用 GRADE 证据等级评价对所有的结果变量的证据质量强度进行分级评估。此外，我们还将考察被试年级、干预周期、干预特征、实验设计对干预效果的影响。

2　研究方法

2.1　文献检索

在英文数据库 PubMed、ISI Web of Science、EMBASE、Science Direct、ProQuest 中，使用关键词组合(Bully*/cyberbully*/anti-bully*) and (school) and (intervention/prevention/program/outcome/evaluation/effect/tackling)进行检索。同时在中国知网、万方和维普数据库中，使用关键词组合(欺凌/反欺凌/网络欺凌/霸凌/欺负 And 学校/校园 And 干预/预防/效果)检索。检索的截止时间为 2018 年 2 月。

2.2　文献纳入排除

对于检索到的文献，根据以下 9 条标准进行筛选和排除。具体的纳入排除流程见图 1。

(1)纳入主题为反欺凌干预项目的效果研究，针对其他暴力行为(例如枪支使用)的项目研究予以排除。

(2)根据欺凌的明确定义。欺凌形式包括口头的、身体的或者心理的攻击，虽然是攻击的一种形式，但是不等完全同于攻击或者校园暴力。

(3)干预项目应用于社区、学校、班级、学生或者家长中，关注的是学生对欺凌行为的态度、欺凌和受到欺凌行为的变化。

(4)随机对照实验或者准实验设计，包含控制组。

(5)结果变量包括学生对欺凌行为的态度，以及欺凌他人或者受欺凌行为结果。结果报告方式包括自评报告、同伴评分、教师或者家长评分及外显的行为观察。

(6)对结果变量的测量时间点至少包括干预前测，干预后测。

(7)研究中必须包含能够计算效果量大小的统计信息，如样本量，均值和标准差，F 值，p 值等。如果缺少该信息，则先尝试联系作者，尽可能获取关键变量的统计值。

(8)审查使用同一个研究项目中同一批数据发表的不同期刊论文，相同的结果不再重复纳入。

(9)所选研究不仅仅包括同行评审期刊论文，还包括学位论文、书的章节等。

2.3　文献编码

详细阅读和编码每一篇纳入元分析的文献，包括以下信息：(1)出版特征：出版年份；(2)干预特征：干预周期；干预措施(a.教育；b.系统干预，包括制定纪律，家长、学生教育等)；(3)干预对象：年龄或者年级；主要对象(a.欺凌受害者；b.欺凌者；c.所有学生) (4)结果测量：结果报告方式(a.自评；b.他评，包括同伴评价，老师评价，实验者观察)。

对于结果变量，我们根据干预项目针对的具体对象进行区分。有的反欺凌干预项目直接针对欺凌者或者欺凌受害者，这种类型的干预直接命中目标，通常有较好的干预效果。而另外一些针对广大学生的干预项目，是普遍型干预，通常有着较低的效果量，但仍然被认为是有效的，因为在普遍干预的学生中发生欺凌行为的比例是相对较低的。已有元分析没有具体划分不同类型的干预项目，可能混淆反欺凌干预效果。因此，我们分

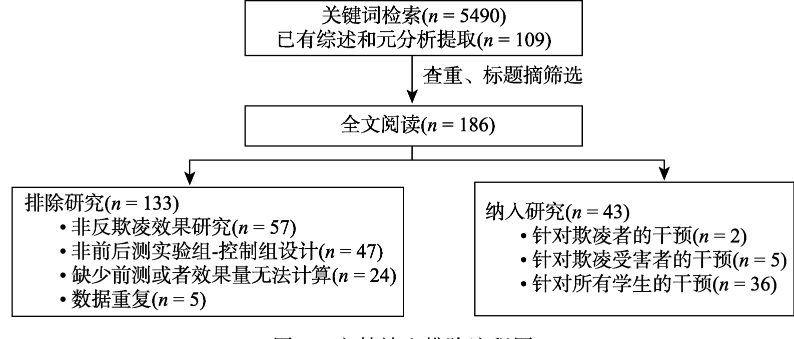


图 1　文献纳入排除流程图

别计算了针对欺凌者的干预项目对其欺凌行为的改善结果；针对欺凌受害者干预项目对其受欺凌及心理健康的效果；针对所有学生的干预项目对欺凌行为和受欺凌行为的干预效果。

2.4 数据分析

首先计算每篇研究每个结果变量的标准化均数差(standard mean difference, SMD)。和以往大部分元分析不同的是，为了减少实验组和控制组基线水平的差异，纳入研究为包含控制组的前后测实验设计(pre-post with control, PPWC) (Hoyt & Del Re, 2017)。计算效果量时考虑了干预前测的数据，先计算两组在结果变量上的前后测变化量的差异，再参照前测数据的标准差进一步标准化(a)，其中 S_{pre} 表示的是前测的合并方差，通过实验组和对照组样本量和标准差求得(b)。

$$d_{PPWC}=\frac{\left(T_{post}-T_{pre}\right)-\left(C_{post}-C_{pre}\right)}{S_{pre}} \quad (a)$$

$$S_{pre}=\left(\frac{n_T-1S_{T1}^2-\left(n_C-1\right)S_{C1}^2}{n_T+n_C-2}\right)^{0.5} \quad (b)$$

当样本量小时，SMD 估计存在偏差，Hedges 建议使用校正后的标准化均数差作为效果量，即 Hedges' g 值(Cooper, Hedges, & Valentine, 2009)。本研究采用CMA 3.0软件对研究的异质性进行检验，计算 Hedges' g，同时进行出版偏差和亚组分析。

异质性检验采用 Q 值和 I^2 评估纳入研究结果的差异性(Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006)。Q 统计量效应量的加权离均平方和，服从自由度为 $k-1$ 的卡方分布，Q 值越大，p 值越小，说明元分析纳入研究的异质性越大。I^2 反应了异质性部分在效应量中总的变异中所占的比重。一般认为 $p < 0.05$ 或者 $I^2 > 50\%$研究间存在显著的异质性。在异质性较大的情况下，建议计算效果量时使用随机效应模型(Hedges & Vevea, 1998)。为了进一步分析效果量的异质性来源，对元分析结果还进行了敏感性分析和亚组分析。敏感性分析采取逐一剔除每个纳入的研究再进行效应量的合并，可以找出效应量异常的研究。亚组分析从临床异质性和方法学异质性的角度来探讨异质性的来源。本研究分析的亚组变量包括干预周期、干预特征、被试年龄、实验设计。

发表偏倚指的是在同类研究中，阳性结果的论文比阴性结果的论文更容易被接受和发表的现象，对元分析结果的推广具有重要影响。本研究采用漏斗图(funnel plots)对发表偏倚进行评估。漏斗图是以效果量为横轴，样本量为纵轴作散点图。理论上认为小样本研究估计的效果量差异较大，分散在漏斗图的底部，大样本的研究估计的效果量的变异程度逐渐降低，因而趋于密集在漏斗图顶部。如果漏斗图不对称或者不完整则说明元分析结果可能存在较大的发表偏倚。由于漏斗图呈现结果较为粗糙，主观性较强，因而进一步采用剪补法(trim and fill)检验结果的稳定性(Duval & Tweedie, 2000)。此外，Egger's 检验也被运用于发表偏倚的检验(郑辉烈，王忠旭，王增珍，2009)。

2.5 GRADE 系统证据质量

GRADE 系统是一套证据评价系统，是当前证据质量和推荐强度分级的国际标准之一(Alonso-Coello et al., 2016)。由于相对于其他证据评价系统更具有代表性，在系统评价、卫生技术评估及指南制作中被广泛运用。GRADE 将证据质量分为四个等级：高，中，低，极低。质量高低显示的是进一步研究是否会影响或者改变该疗效的评估结果。

在 GRADE 中，基于 RCT 得出的证据被定为高质量，非 RCT 研究属于低质量证据。但是二者证据质量还要结合具体的标准进行调整。其中，可能降低证据质量的因素共有 4 个：(1)研究的局限性；(2)结果的不一致；(3)间接证据；(4)结果的精确度；(5)发表偏倚。可能提高证据质量的因素包括：(1)大效应值；(2)可能的混杂因素会降低疗效；(3)存在剂量−效应关系。

根据这套定量的证据分级系统，采用 GRADEpro 3.6 软件对本研究中结果变量的 RCT 研究证据及非 RCT 研究证据结果进行评估，呈现元分析结果的证据质量。

3 结果

3.1 纳入研究基本信息

共纳入 43 个反欺凌干预效果研究，有效样本 51 个，总样本量 41363 人(实验组 $N = 23582$，控制组 $N = 17781$)。其中，中文文献 4 篇，英文文献 39 篇。针对欺凌者的干预样本 4 个，随机对照样本 2 个；针对欺凌受害者干预样本 6 个，随机对照样本 3 个；针对所有学生干预样本 42 个，随机对照样本 20 个。学生年龄 6~16 岁之间，包括小学生、初中生和高中生。详细信息见表 1。

表 1　纳入文献基本信息

研究	干预实施时间	样本	干预对象	干预周期	年龄	年级	地点	实验设计	结果报告	干预项目简要
Albayrak, Yıldız, & Erol, 2016	2011/03~2011/05	367	学生	2 个月	12~15	elementary (6~8)	Istanbul	non-RCT	自评	包含多水平的干预策略：(1)父母教育；(2)学生的态度干预、知识教育和技能培养；(3)公共媒体信息教育；(4)提供针对家长或者学校如何处理欺凌行为的建议。
Andreou, Didaskalou, & Vlachou, 2007	2003/12~2004/05	454	学生	4 周	10.23	primary (4~6)	Greece	non-RCT	自评	基于动力关系理论，形式包含课程和课外活动。主要出发点：(1)提高学生的意识和认识；2.促进学生自省；3.制定班级规则。
Baldry & Farrington, 2004	—	237	学生	3 周	10~16	middle/high	Italy	RCT	自评	视频干预：3 个视频和一本指导手册，包含三个部分：(1)青少年间的欺凌；(2)看见家庭暴力的孩子们；(3)暴力循环；
Battey, 2008	—	249	学生	1 学期	12.6	secondary (7)	USA	non-RCT	自评	BPCCC，课程和技巧
Bauer, Lozano, & Rivara, 2007	2003~2005	6518	学生	2 学年	—	middle (6~8)	USA	non-RCT	自评	Olweus Bullying Prevention Program:(1)学校水平：校规，欺凌委员会，管理人员及教师培训；(2)课堂讨论，加强对校规的实行；(3)对欺凌和被欺凌人员的干预；(4)社区管理。
Beran, Tutty, & Steinrath, 2004	—	102	学生	3 个月	—	elementary (4~6)	Alberta	non-RCT	自评	Dare to Care: Bully Proofing Your School.(1)提供专业的培训；(2)学校，家长和学生共同制定纪律；(3)课堂教育：知识、态度和技能。
Berry & Hunt, 2009	2004	46	欺凌受害者	8 周	13.04	secondary (7~10)	Australia	RCT	自评/父母评价	课程干预：基于认知行为，集中于内化行为，提高学生自尊，教授社交技巧和应对方式。
Boulton & Flemington, 1996	—	170	学生	2 周	11~14	semi-rural secondary (7~10)	England	RCT	自评	视频干预：Sticks and Stones，包含几个欺凌场景，引起同情和害怕。
Brown, Low, Smith, & Haggerty, 2011	2008/12~2009/05	2490	学生	1 年	8.9	3~6	USA	RCT	教师评价	Steps to Respect (STR)，提高学生社交能力，降低同伴拒绝，提高旁观者的积极行为，同伴对欺凌的干预，增强学校教育，包含(1)班级课程：友谊、社交技能、情绪管理、求助、共情培养、对欺凌的意识等；(2)学校：反欺凌规定和政策；培训和报告；提高意识，加强管理。
Couch, 2015	2013~2014	206	学生	1 学年	8.97	3~4	USA	RCT	自评/教师评价	BLP，指定阅读欺凌书，进行讨论，学习应对技能

续表 1

研究	干预实施时间	样本	干预对象	干预周期	年龄	年级	地点	实验设计	结果报告	干预项目简要
DeRosier, 2004	—	381	学生	8 周	8.6	primary	USA	RCT	自评	Social Skills GRoup INtervention (S.S.GRIN):目的是建立学生社交技能，强化态度，处理问题的技巧，有指导手册。包括社交反应训练，角色扮演，模仿，行动等。
Doğan et al., 2017	2015/09~2016/06	411	学生	1 年	10.06	5	Turkey	non-RCT	自评	Social Competence Program, ViSC:社会生态反欺凌项目，包含了学校，班级和个体水平。学校水平：训练教师；班级水平：训练反欺凌的能力，责任感，情绪觉察和处理。个体水平：谈话。
Domino, 2011	2009~2010	336	学生	1 学期	12~18	7	USA	non-RCT	自评	LEAD, 16 单元的课程学习
Elledge, Cavell, Ogle, & Newgent, 2010	—	24	欺凌受害者	1 学期	11.06	elementary (4~5)	USA	non-RCT	自评/同伴/教师评价	Lunch Buddy Mentoring:一个被试配对一个受过训练的大学生在午餐时间共处；陪伴人员经过训练，每周进行讨论。
Fekkes, Pijpers, & Verloove-Vanhorick, 2006	1999/11~2000/05	70	学生	1 学年	9~12	elementary	Netherlan	RCT	自评	Bullying in School: How to Deal With It:包含校规，课程等等
Fox & Boulton, 2003	1999/12~2000/03	28	欺凌受害者	8 周	9~11 (9.6)	primary	USA	non-RCT	自评	Social Skills Training (SST) Programme:教授如何使用社交问题解决技巧，如何使用放松技巧，如何积极思考，如何调整行为，如何使用口头技巧。
Fraser, 2004		112	受欺凌者	5 周		secondary (6~8)	USA	non-RCT	自评	Bully Proofing Your School (BPYS)课程
Frey et al., 2005	—	2252	学生	1 年	—	primary (3~6)	USA	RCT	自评	Steps to Respect program：目的是减少欺凌问题，a.提高意识和责任感; b.增强社会责任感; c.教授社交技能。包含(1)训练：咨询师，管理者和老师。(2)课堂课程；知识和技能学习; (3)父母参与。
Gradinger, Yanagida, Strohmeier, & Spiel, 2015	2009/09~2010/06	1639	学生	1 学年	11.7	secondary	Australia	RCT	自评	ViSC
Hoglund, Hosan, & Leadbeater, 2012	2000~2006	432	学生	6 年	6.3	primary (1~3)	Canada	non-RCT	自评	Using your wits：教育类，远离欺凌和求助。
Hunt, 2007	2001~2002	400	学生	1 年	12~15	secondary	Australia	RCT	自评	education anti-bullying：家长和教师会议，课堂课程，提高意识，对欺凌的识别，对被欺凌者的共情，反欺凌的技巧。

续表 1

研究	干预实施时间	样本	干预对象	干预周期	年龄	年级	地点	实验设计	结果报告	干预项目简要
Joronen, Konu, Rankin, & ÅstedtKurki, 2011	—	116	学生	1 学年	10.4	4~5	USA	non-RCT	自评	课程教育：戏剧项目，包括课堂环节，家庭活动，家长会谈。
Kärnä et al., 2011; Salmivalli, Kärnä, & Poskiparta, 2011	2007/05~2007/12	8166	学生	1 学期	10~12	4~6	Finland	RCT	自评	KiVa antibullying program：大型项目，提高意识，对欺凌受害者的同情，学会社交技巧。
Kärnä et al., 2013	2008/05~2009/05	10382	学生	1 年	—	primary (1~3)/ secondary (7~9)	Finland	RCT	自评	KiVa
Krueger, 2010	—	47	学生	2 周	—	elementary	USA	RCT	他评	Take a Stand, Lend a Hand: Stop Bullying Now!课程
McLaughlin, 2009	—	68	学生	4 周	11.35	6	USA	non-RCT	自评	CBT 课程
Menesini, Codecasa, Benelli, & Cowie, 2003	1998/10~1999/05	293	学生	1 年	10~14	middle (6~8)	Italia	non-RCT	自评	阶段 1：课堂干预，提高意识；阶段 2：同伴支持；阶段 3：训练倾听和交流技巧；阶段 4：类似团体辅导；阶段 5：推广.
Meyer & Lesch, 2000	—	12	欺凌者	10 周	12~16	primary	South Africa	RCT	自评/同伴评价	强化积极行为，自我观察，家庭任务，模仿，角色扮演等。
Naidoo, Satorius, de Vries, & Taylor, 2016	2013/02~2013/10	499	学生	4 个月	16.9	secondary	South Africa	RCT	自评	课程教育
Stan & Beldean, 2014	—	232	学生	—	—	middle (5~8)	罗马尼亚	non-RCT	自评	Program Achieve. You Can Do It!：课程指导，情绪管理和技能训练
Nocentini & Menesini, 2016	2013/05~	905	学生	1 学年	—	primary (4~6)	Italia	RCT	自评	KiVa
Palladino, Nocentini, & Menesini, 2016	2011~2012	450	学生	1 学年	14.91	9	Italia	non-RCT	自评	NoTrap! Program
van der Ploeg, Steglich, & Veenstra, 2016	2012/10~2013/05	134	受害者	1 学年	9.15	2~6	Dutch	non-RCT	自评	Kiva
Rawana, Norwood, & Whitley, 2011	—	103	学生	8 个月	11	4~8	Canada	non-RCT	自评	Strengths in Motion Bullying Prevention Program
Şahin, 2012	2009~2010	36	欺凌者	11 周		primary (6)	Turkey	RCT	自评	共情训练

续表 1

研究	干预实施时间	样本	干预对象	干预周期	年龄	年级	地点	实验设计	结果报告	干预项目简要
Solomontos-Kountouri, Gradinger, Yanagida, & Strohmeier, 2016	2012/10~2013/04	343	学生	1 学年	12.5	7	Cyprus	non-RCT	自评	ViSC
Stevens, Bourdeaudhuij, & van Oost, 2000	1995/10~1997/05	550	学生	2 学年	10~16	primary/secondary	Flemish schools	RCT	自评	The Flemish school-based anti-bullying intervention programme：基于 Olweus 项目，包含三个模块，有视频和指导手册。(1)干预学校环境：制定政策，提高意识。(2)青少年的课程学习，提高认知，制定班级规则，教授问题解决策略，社交技能训练。(3)针对欺凌者或者受害者
Trip et al., 2015	2011/10~2012/05	431	学生	1 学年	11.82	6	Romalia	RCT	自评	ViSC
Yanagida, Strohmeier, & Spiel, 2016	2009/09~2010/06	431	学生	1 学年	11.7	secondary	Vienna	non-RCT	自评	ViSC
高妍, 2011	2010/10~2010/12	62	学生	8 周	—	primary	China	non-RCT	自评	团体辅导
任丽娜, 2007	—	315	学生	—	—	secondary	China	non-RCT	自评	多水平干预：教师及家长会议；班级活动；个别辅导
杨婉秋, 2004	2003/09~1004/04	115	学生	7 周	—	primary/secondary	China	non-RCT	自评	团体心理咨询：班级和个体水平上的干预
张文新, 鞠玉翠, 2009	—	354	学生	5 周	—	primary	China	non-RCT	自评	班级水平的干预，包含(1)科学研究的程序和方法；(2)关于欺负课题的有关知识；(3)行动研究法；(4)具体干预策略—头脑风暴法、自信心训练、角色扮演技术等。

3.2 反欺凌项目干预效果

异质性检验结果显示, 反欺凌干预研究间差异性大。对于针对欺凌者的干预样本, 欺凌行为 Q = 48.23 ($p < 0.001$), $I^2 = 93.79$。对于针对欺凌受害者的干预样本, 受欺凌行为 Q = 18.45 ($p < 0.01$), I^2 = 73.04, 心理健康 Q = 7.95 ($p > 0.05$), I^2 = 49.66 (见表 2)。针对所有学生的反欺凌干预样本中, 态度干预研究的 Q = 131.79 ($p < 0.001$), I^2 = 93.05; 行为干预研究 Q = 330.23 ($p< 0.001$), I^2 = 87.58; 其中欺凌行为改变的 Q = 123.96 ($p < 0.001$), I^2 = 70.15; 受欺凌状况改变的 Q = 354.73 ($p < 0.001$), I^2 = 89.57 (见表 3)。敏感性分析结果显示, 不存在效果异常研究。

针对欺凌者干预项目对其欺凌行为的干预效果 g = 0.57 (95%CI: 0.08, 1.05; Z = 2.30, $p < 0.05$); 针对欺凌受害者干预项目对其受欺凌行为干预效果 g = 0.42 (95%CI: 0.10, 0.74; Z = 2.30, $p < 0.05$), 对其心理健康干预效果 g = 0.40 (95%CI: 0.14, 0.65; Z = 3.04, $p < 0.01$) (见表 2)。对于所有学生的干预项目, 对态度的影响效果 g = 0.40 (95%CI: 0.19, 0.61; Z = 3.79, $p < 0.001$), 对行为的改变效果 g = 0.21 (95%CI: 0.15, 0.26, Z = 7.09, $p < 0.001$), 其中对于欺凌行为的效果 g = 0.17 (95%CI: 0.12, 0.22; Z = 6.91, $p < 0.001$), 对于受欺凌行为的效果 g = 0.19 (95%CI: 0.12, 0.26; Z = 5.27, $p < 0.001$) (见表 4)。

3.3 发表偏倚

漏斗图 2~图 4 分别检验了针对所有学生的反欺凌干预项目对态度、欺凌行为及受欺凌行为的干预效果研究是否存在发表偏倚。针对所有学生的干预研究显示, 学生态度改变漏斗图经过剪补法校正后, 新增研究为 0, 结果较为稳健。以 0.05 为检验水准, Egger’s 检验(p = 0.100)提示本次纳入研究的发表偏倚较小, 和漏斗图及剪补法分析结果一致。但是, 不管是欺凌行为还是受欺凌行为的漏斗图都不对称, 欺凌行为效果经剪补法校正后, 新增研究 14 个, 校正后随机效应 g = 0.08 (95%CI: 0.03, 0.14); 受欺凌行为效果经剪补法校正后, 新增研究 13 个, 校正后随机效应 g = 0.05 (95%CI: −0.02, 0.12), 说明二者可能存在出版偏差。Egger’s 检验($p_{欺凌}$ = 0.000, $p_{受欺凌}$ = 0.003)提示本次纳入计算欺凌行为和受欺凌行为改变效果的研究存在发表偏倚, 和漏斗图及剪补法分析结果一致。

3.4 亚组分析

如表 4 所示, 亚组分析结果显示, 对欺凌行为改变效果量, 干预周期, 干预特征是显著影响效果量的因素。干预周期小于 1 学期的效果量(g = 0.26, $p < 0.001$)高于干预周期大于 1 学期的效果量(g = 0, 11, $p < 0.001$), 具有统计学意义($p < 0.05$)。多水平干预效果量(g = 0.30, $p < 0.001$)高于单独个体水平干预(g = 0.10, $p < 0.001$), 差异统计学上显著($p < 0.01$)。非随机研究效果量(g = 0, 28, $p < 0.001$)高于随机对照研究效果量(g = 0, 12, $p < 0.01$), 差异显著($p < 0.05$)。同样地, 如表 5 所示, 对于受欺凌行为, 干预周期和干预特征为显著地

表 2　针对欺凌者或欺凌受害者的干预效果

干预对象	结果	N	G	95%CI	Z	Q	I^2
欺凌者	欺凌行为	4	0.57	0.08, 1.05	2.30*	48.23***	93.79
欺凌受害者	受欺凌行为	6	0.42	0.10, 0.74	2.60**	18.55**	73.04
	心理健康	5	0.40	0.14, 0.65	3.04**	7.95	49.66

注：*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$。下同

表 3　针对所有学生的反欺凌项目预防干预效果

结果	N	G	95%CI	Z	Q	I^2
态度	11	0.40	0.19, 0.61	3.79***	131.79***	94.41
行为	42	0.21	0.15, 0.26	7.09***	330.23***	87.58
欺凌	38	0.17	0.12, 0.22	6.91***	123.96***	70.15
受欺凌	38	0.19	0.12, 0.26	5.27***	354.73***	89.57

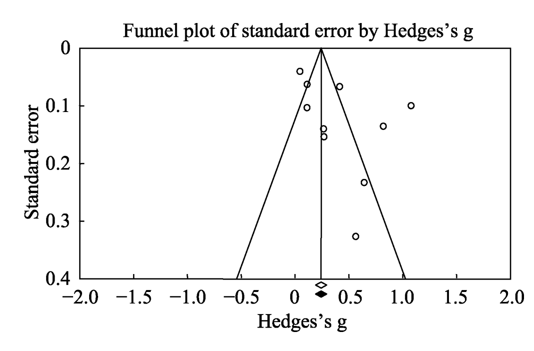


图 2 针对所有学生的态度干预效果漏斗图; 剪补研究 = 0

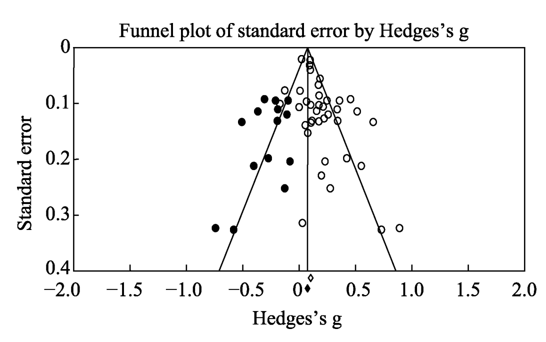


图 3 针对所有学生的欺凌行为干预效果漏斗图; 剪补研究 = 14, 随机效应 g = 0.08, 95%CI: 0.03, 0.14

调节变量。干预周期小于 1 学期的效果量(g = 0.31, $p < 0.001$)高于干预周期大于 1 学期的效果量(g = 0.10, $p < 0.001$), 具有统计学意义($p < 0.01$)。单独个体水平干预效果量(g = 0.10, $p < 0.001$)低于多水平干预(g = 0.39, $p < 0.001$), 差异统计学上显著(p = 0.001)。非随机研究效果量(g = 0, 28, $p < 0.001$)高于随机对照研究效果量(g = 0, 12, $p < 0.01$), 差异显著($p < 0.05$)。初高中被试的干预效果高于小学生的干预效果, 但是差异不具有统计学上的显著意义。

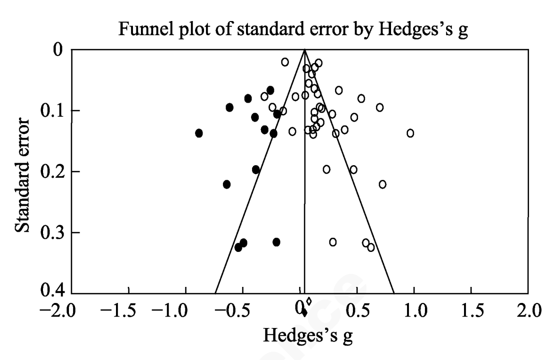


图 4 针对所有学生的受欺凌行为干预效果漏斗图; 剪补研究 = 13, 随机效应 g = 0.05, 95%CI: −0.02, 0.12

3.5 GRADE 证据质量及推荐等级

根据 GRADE 系统对系统评价的证据质量进行评估, 结果显示, 对于欺凌者的干预效果的证据质量为极低。针对欺凌受害者的干预效果, 不管是 RCT 证据还是非 RCT 研究证据质量都为低。针对所有学生的干预项目, 欺凌行为或者受欺凌行为 RCT 证据为低质量, 非 RCT 研究为极低质量, 不过态度改变 RCT 研究的证据质量为中等(表 6)。

表 4 欺凌干预对欺凌行为改变效果量的亚组分析

亚组变量	N	G	95%CI	Z	Q	I^2	p
年级							0.228
小学	17	0.14	0.09, 0.18	6.18***	21.37	25.14	
初高中	21	0.20	0.11, 0.29	4.27***	100.77***	80.15	
干预周期							0.004
< = 1 学期	20	0.26	0.18, 0.35	5.96***	64.08***	70.35	
> 1 学期	18	0.11	0.05, 0.17	3.69***	45.60***	62.72	
干预特征							0.001
多水平	22	0.30	0.20, 0.40	5.97***	26.89*	44.22	
个体水平	16	0.10	0.07, 0.17	4.73***	62.90***	66.61	
实验设计							0.016
随机	19	0.12	0.08, 0.17	5.08***	47.99***	62.49	
非随机	19	0.28	0.16, 0.39	4.67**	60.86**	70.43	

表 5　欺凌干预对受欺凌行为改变效果量的亚组分析

亚组变量	*N*	*G*	95%CI	*Z*	*Q*	I^2	*p*
年级							0.457
小学	18	0.21	0.14, 0.28	5.93***	63.73***	73.33	
初高中	20	0.16	0.04, 0.28	2.72*	234.53***	91.90	
干预周期							0.003
<=1 学期	17	0.31	0.20, 0.42	5.59***	98.35***	83.73	
>1 学期	21	0.10	0.02, 0.18	2.38*	170.90***	88.30	
干预特征							0.001
多水平	24	0.39	0.24, 0.55	4.88***	52.19***	77.09	
个体水平	14	0.10	0.03, 0.18	2.78*	232.15***	90.12	
实验设计							0.042
非随机	20	0.28	0.15, 0.41	4.29***	130.97***	85.49	
随机	18	0.12	0.04, 0.21	2.77**	198.72**	91.45	

表 6　结果变量证据 GRADE 总结表

干预对象	结局指标	实验设计	纳入研究	结果的不一致	间接证据	结果不精确	发表偏倚	实验组	对照组	相对效应	证据质量	结果变量重要性
欺凌者	欺凌行为	RCT[1]	2	严重[2]	无	严重[3-4]	无	24	24	0.96(−1.04, 2.95)	⊕OOO	关键
受凌者	受欺凌行为	RCT[1]	2	严重[2]	无	无	无	209	216	0.79(−0.56, 2.14)	⊕⊕OO	关键
		非 RCT	3	无	无	无	无	104	63	0.33(0.07, 0.59)	⊕⊕OO	关键
	心理健康	RCT[1]	2	严重[2]	无	无	无	209	216	0.57(−0.13, 1.27)	⊕⊕OO	关键
		非 RCT	2	无	无	无	无	93	49	0.49(0.15, 0.84)	⊕⊕OO	关键
所有学生	欺凌行为	RCT[1]	14	严重[2]	无	严重[3]	无	1110	1101	0.12(0.08, 0.17)	⊕⊕OO	关键
		非 RCT	10	严重[2]	无	严重[3]	严重	1505	1776	0.28(0.20, 0.40)	⊕OOO	关键
	受欺凌行为	RCT[1]	13	严重[2]	无	严重[3]	无	12630	10731	0.12(0.04, 0.21)	⊕⊕OO	关键
		非 RCT	12	严重[2]	无	严重[3]	严重	7697	3791	0.28(0.15, 0.41)	⊕OOO	关键
	态度	RCT[1]	6	无	无	无	无	1479	1589	0.34(0.06, 0.61)	⊕⊕⊕O	重要

注：1 盲法和分配隐藏缺失; 2 干预措施不同、被试特征差异导致结果不一致; 3 单一的问卷自评测量; 4 被试数量小

4　讨论

本研究更新了已有的元分析，对反欺凌项目的干预效果研究进行了系统分析，结果发现欺凌干预是有效的，不同的反欺凌项目效果有所差异。亚组分析显示干预周期越短，干预水平越多，反欺凌干预效果越好。不过进一步的证据质量评价显示现有的研究证据质量较差。

直接针对欺凌者或者欺凌者的干预项目具有中等效果量。对于欺凌者的干预包括共情训练、行为矫正，目的在于减少欺凌者的攻击性和欺凌行为(Garandeau, Vartio, Poskiparta, & Salmivalli, 2016; Şahin, 2012)。对欺凌受害者的干预包括社交技能训练、成年人监管、提供社会支持等，目的在于减少欺凌带来的负面心理问题，并且提高个体的社交能力、自信心、求助能力等以降低受到欺凌的频率(Elledge et al., 2010; Fox & Boulton, 2003; van der Ploeg et al., 2016)。尽管研究显示这些干预项目是有效的，但是由于研究数量有限，元分析结果可靠性较低。可能的影响干预有效性的原因有以下几个方面：一方面，及时、准确地识别欺凌存在一定的困难。由于许多欺凌事件是

没有被报告的, 因此学校或者教师未能够及时、准确地识别欺凌者和欺凌受害者(Hymel & Swearer, 2015)。另一方面, 这些干预是事后干预, 管理者更加希望能够预防欺凌行为的发生(Zych, Ortega-Ruiz, & Del Rey, 2015)。还有一种观点认为, 欺凌是群体行为, 旁观者或者欺凌行为见证者的态度和行为也至关重要(Forsberg et al., 2018)。

因此, 针对所有学生的干预项目融合了预防和干预, 面对的对象不只是欺凌者或者受欺凌者, 还包括旁观者、家长、教师等, 干预内容包括知识教育、态度影响、纪律、家长参与等, 是更为常见的反欺凌项目(Evans et al., 2014)。元分析显示, (1)干预对学生态度有中等的效果量, 对于欺凌或者受欺凌行为的效果量小但是显著。(2)态度改变效果量高于行为改变效果量。相对于行为, 态度更容易发生转变(Maio & Haddock, 2014)。另外, 对于一部分学生而言, 欺凌或者受欺凌行为的基线水平本身就较低, 因此干预后的变化较小。Jiménez-Barbero 等人(2016)的结果和本研究相对一致, 但是效果量更低, 可能的原因在于本研究计算效果量的时候考虑了基线数据, 根据 Carlson 和 Schmidt (1999)的检验, 只用后测数据计算的效果量可能会低估真正的效果。(3)被试的欺凌行为和受欺凌行为改变效果量没有差别。Merrell 等(2008)发现不同测量工具下, 报告的欺凌结果和受欺凌结果并不相同。尽管本研究同时纳入了自评、教师评价、家长评价、观察数据, 但是大部分研究采用的是自评测量, 因而没有对测量方式或者工具进一步细分, 这可能弱化了欺凌改变效果和受欺凌改变效果之间的差异。

对所有学生的干预研究异质性大, 进一步的亚组分析显示, 干预周期和干预特征是欺凌/受欺凌行为改变量的显著调节变量。和一个已有的元分析一样, 干预周期越短, 干预效果越好, 这似乎让人有些匪夷所思(Jiménez-Barbero et al., 2016)。有一些研究指出, 反欺凌干预也存在量效反应(dose-response), 表现为越密集, 强度越高的干预项目, 效果越好(Ttofi & Farrington, 2011)。可能对于那些短周期的研究, 干预强度更高, 更密集。此外, 对于那些干预周期很长的研究, 可能在干预初期, 起到的教育或者威慑效果更好, 但是, 随着学生的适应, 对欺凌的关注慢慢降低, 导致欺凌事件重新发生。

系统多水平的干预项目的效果量高于单水平的教育干预。从社会生态视角来看, 欺凌不仅仅是“欺凌者”和“受害者”之间的事, 除了个体的人格特质、行为, 家庭、同伴、学校等系统都与欺凌行为的发生密切相关(Hong, Lee, Lee, Lee, & Garbarino, 2014)。多数的教育干预针对的是个体的知识、态度和行为, 只在个体水平上进行干预。越来越多的干预项目不仅仅从个体水平出发, 还延伸到学校、家庭、社区甚至国家政策中。例如芬兰的 Kiva 反欺凌项目不仅仅向学生提供反欺凌课程, 还向教师提供反欺凌设备材料, 改变旁观者的角色, 给家长提供反欺凌指导, 设置“反欺凌街道”论坛, 积极应对网络欺凌等等, 从不同的水平进行干预(Yang & Salmivalli, 2015)。有的干预项目还借助公共媒体传播反欺凌信息(Albayrak et al., 2016)。不过, 多水平的干预相对地需要投入更多的人力、物力和财力, 但是从效果上来说, 确实有助于欺凌的预防和干预。

随机实验设计和非随机实验设计下的反欺凌项目干预效果也有显著差异。已有研究指出非随机实验设计或者队列设计的效果量会高于严谨的实验设计, 但是无法提供解释(Ttofi & Farrington, 2011)。在控制了基线水平的差异后, 本研究发现非随机对照干预的效果量更高。可能的原因是非随机设计的研究两组被试可能是不同质的, 在没有使用随机化平衡异质性的情况下, 容易获得“好效果” (Hoyt & Del Re, 2017)。

小学生还是中学生从干预项目中获益更多? 本研究显示二者没有显著差异, 与本研究结果不同的是, Lee, Kim 和 Kim (2015)对 13 个校园欺凌项目的元分析中发现对初中生的干预效果优于小学生, 因为初中生的认知发展更为成熟, 更容易接受干预。而另一些研究又指出对小学生的干预效果更好, 因为他们可塑性更强, 也更会听老师的话(Jiménez-Barbero et al., 2016)。包括本研究在内的这些元分析的不一致可能在于过于笼统地将学生划分为小学生或者初中生。Yeager, Fong, Lee 和 Espelage (2015)认为这种研究间的亚组分析(between-study)是有偏差的, 他们对 19 个反欺凌项目研究采用了多水平的元分析, 其中年龄作为第一个水平, 结果发现干预效果实际上随着年龄的上升而下降。

值得引起重视的是, 根据 GRADE 证据分级

系统对各个结果变量的质量评价结果显示，不管是来自随机对照实验还是非随机的研究证据质量多数低甚至极低，只有态度改变的 RCT 证据质量为中等。尽管心理学的干预和医学研究还是存在一定的差异，特别是对于大规模学校干预，随机对照已经是相对困难，更不用说双盲和隐藏分配等设计，但是研究者应该致力于尽可能减少干扰因素，详细说明实验过程，并尽可能标准化实验设计。总之，在反欺凌干预领域，还需要更多严谨的研究对其效果进行检验。

5　局限和展望

本研究存在几个局限：(1)可能存在“灰色”文献检索遗漏；(2)由于测量方式主要以自评为主，没有对结果变量进行细分，而测量工具可能是影响效果量的因素之一；(3)干预研究的异质性很大，尽管亚组分析找到了两个影响效果量的变量，但是可能存在更多的影响因素有待进一步进行检验。

从研究结果来看，未来的研究需要进一步考察具体哪些要素影响干预项目的效果，同时尽可能提升实验设计的质量，将干预项目应用到更多的地区、人群中，并根据实际情况进行调整。并且，在设计干预项目时，尽可能从多个水平出发，更加全面和系统地对预防和干预欺凌。此外，我国校园欺凌也经常发生(张文新，2002)，越来越多的研究者关注到校园欺凌的不良后果和风险因素(雷雳，王燕，郭伯良，张雷，2004)，但是干预研究还相对较少，未来可进一步结合已有的研究结果，设计符合我国国情的反欺凌项目，并进行效果检验。

参考文献

(*为纳入元分析文献)

*高妍. (2011). *小学生欺负行为干预研究*(博士学位论文). 山西大学.

雷雳，王燕，郭伯良，张雷. (2004). 班级行为规范对个体行为与受欺负关系影响的多层分析. *心理学报, 36*(5), 563–567.

*任丽娜. (2007). *初中生受欺负者的干预研究*(硕士学位论文). 山西大学.

*杨婉秋. (2004). *团体心理咨询降低小学生欺负行为的实验研究*(硕士学位论文). 云南师范大学.

郑辉烈，王忠旭，王增珍. (2009). Meta 分析中发表偏倚的 Begg's 检验、Egger's 检验及 Macaskill's 检验的 SAS 程序实现. *中国循证医学杂志, 9*(8), 910–916.

张文新. (2002). 中小学生欺负/受欺负的普遍性与基本特点. *心理学报, 34*(4), 387–394.

*张文新，鞠玉翠. (2009). 小学生欺负问题的干预研究. *中国教育学前沿,* (1), 95–99.

*Albayrak, S., Yıldız, A., & Erol, S. (2016). Assessing the effect of school bullying prevention programs on reducing bullying. *Children and Youth Services Review, 63*, 1–9.

Alonso-Coello, P., Schünemann, H. J., Moberg, J., Brignardello-Petersen, R., Akl, E. A., Davoli, M., ... Oxman, A. D. (2016). GRADE evidence to decision (EtD) frameworks: A systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ, 353*, i2016.

*Andreou, E., Didaskalou, E., & Vlachou, A. (2007). Evaluating the effectiveness of a curriculum-based anti-bullying intervention program in Greek primary schools. *Educational Psychology, 27*(5), 693–711.

*Baldry, A. C., & Farrington, D. P. (2004). Evaluation of an intervention program for the reduction of bullying and victimization in schools. *Aggressive Behavior, 30*(1), 1–15.

*Battey, G. J. L. (2008). *Can bullies become buddies? Evaluation of and theoretical support for an experiential education bully prevention curriculum with seventh grade students* (Unpublished doctoral dissertation). Oregon State University.

*Bauer, N. S., Lozano, P., & Rivara, F. P. (2007). The effectiveness of the Olweus Bullying Prevention Program in public middle schools: A controlled trial. *Journal of Adolescent Health, 40*(3), 266–274.

*Beran, T. N., Tutty, L., & Steinrath, G. (2004). An evaluation of a bullying prevention program for elementary schools. *Canadian Journal of School Psychology, 19*(1-2), 99–116.

*Berry, K., & Hunt, C. J. (2009). Evaluation of an intervention program for anxious adolescent boys who are bullied at school. *Journal of Adolescent Health, 45*(4), 376–382.

*Boulton, M. J., & Flemington, I. (1996). The effects of a short video intervention on secondary school pupils' involvement in definitions of and attitudes towards bullying. *School Psychology International, 17*(4), 331–345.

*Brown, E. C., Low, S., Smith, B. H., & Haggerty, K. P. (2011). Outcomes from a school-randomized controlled trial of steps to respect: A bullying prevention program. *School Psychology Review, 40*(3), 423–443.

Carlson, K. D., & Schmidt, F. L. (1999). Impact of experimental design on effect size: Findings from the research literature on training. *Journal of Applied Psychology, 84*(6), 851–862.

Chan, H. C. O., & Wong, D. S. W. (2015). Traditional school bullying and cyberbullying in Chinese societies: Prevalence and a review of the whole-school intervention approach. *Aggression and Violent Behavior, 23*, 98–108.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis, second*

edition. Russell Sage Foundation.

*Couch, L. (2015). *The bullying literature project: An evaluation of a class-wide bullying intervention program* (Unpublished doctoral dissertation). University of California, Riverside.

*DeRosier, M. E. (2004). Building relationships and combating bullying: Effectiveness of a school-based social skills group intervention. *Journal of Clinical Child and Adolescent Psychology, 33*(1), 196–201.

*Doğan, A., Keser, E., Şen, Z., Yanagida, T., Gradinger, P., & Strohmeier, D. (2017). Evidence based bullying prevention in turkey: Implementation of the ViSC social competence Program. *International Journal of Developmental Science, 11*(3-4), 93–108.

*Domino, M. B. (2011). *The impact of take the LEAD on school bullying among middle school youth* (Unpublished doctoral dissertation). Walden University.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463.

*Elledge, L. C., Cavell, T. A., Ogle, N. T., & Newgent, R. A. (2010). School-based mentoring as selective prevention for bullied children: A preliminary test. *The Journal of Primary Prevention, 31*(3), 171–187.

Evans, C. B. R., Fraser, M. W., & Cotter, K. L. (2014). The effectiveness of school-based bullying prevention programs: A systematic review. *Aggression and Violent Behavior, 19*(5), 532–544.

Farrington, D. P., Gaffney, H., Lösel, F. A., & Ttofi, M. M. (2016). Systematic reviews of the effectiveness of developmental prevention programs in reducing delinquency, aggression, and bullying. *Aggression and Violent Behavior, 33*, 91–106.

*Fekkes, M., Pijpers, F. I., & Verloove-Vanhorick, S. P. (2006). Effects of antibullying school program on bullying and health complaints. *Archives of Pediatrics & Adolescent Medicine, 160*(6), 638–644.

Forsberg, C., Wood, L., Smith, J., Varjas, K., Meyers, J., Jungert, T., & Thornberg, R. (2018). Students' views of factors affecting their bystander behaviors in response to school bullying: A cross-collaborative conceptual qualitative analysis. *Research Papers in Education, 33*(1), 127–142.

*Fox, C., & Boulton, M. (2003). Evaluating the effectiveness of a social skills training (SST) programme for victims of bullying. *Educational Research, 45*(3), 231–247.

*Fraser, C. M. (2004). *Bully proofing your school: The effectiveness of a school-wide prevention program with middle school students* (Unpublished doctoral dissertation). Capella University.

*Frey, K. S., Hirschstein, M. K., Snell, J. L., Edstrom, L. V. S., MacKenzie, E. P., & Broderick, C. J. (2005). Reducing playground bullying and supporting beliefs: An experimental trial of the steps to respect program. *Developmental Psychology, 41*(3), 479–490.

Garandeau, C. F., Vartio, A., Poskiparta, E., & Salmivalli, C. (2016). School bullies' intention to change behavior following teacher interventions: Effects of empathy arousal, condemning of bullying, and blaming of the perpetrator. *Prevention Science, 17*(8), 1034–1043.

*Gradinger, P., Yanagida, T., Strohmeier, D., & Spiel, C. (2015). Prevention of cyberbullying and cyber victimization: Evaluation of the ViSC social competence program. *Journal of School Violence, 14*(1), 87–110.

Havik, T., Bru, E., & Ertesvåg, S. K. (2015). School factors associated with school refusal-and truancy-related reasons for school non-attendance. *Social Psychology of Education, 18*(2), 221–240.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods, 3*(4), 486–504.

*Hoglund, W. L., Hosan, N. E., & Leadbeater, B. J. (2012). Using your WITS: A 6-year follow-up of a peer victimization prevention program. *School Psychology Review, 41*(2), 193–214.

Hong, J. S., Lee, C-H., Lee, J., Lee, N. Y., & Garbarino, J. (2014). A review of bullying prevention and intervention in south korean schools: An application of the social-ecological framework. *Child Psychiatry & Human Development, 45*(4), 433–442.

Hoyt, W. T., & Del Re, A. C. (2017). Effect size calculation in meta-analyses of psychotherapy outcome research. *Psychotherapy Research*, *28*(3), 1–10.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods*, *11*(2), 193–206.

*Hunt, C. (2007). The effect of an education program on attitudes and beliefs about bullying and bullying behaviour in junior secondary school students. *Child and Adolescent Mental Health, 12*(1), 21–26.

Hymel, S., & Swearer, S. M. (2015). Four decades of research on school bullying: An introduction. *American Psychologist, 70*(4), 293–299.

Jiménez-Barbero, J. A., Ruiz-Hernández, J. A., Llor-Zaragoza, L., Pérez-García, M., & Llor-Esteban, B. (2016). Effectiveness of anti-bullying school programs: A meta-analysis. *Children and Youth Services Review, 61*, 165–175.

*Joronen, K., Konu, A., Rankin, H. S., & ÅstedtKurki, P. (2011). An evaluation of a drama program to enhance social relationships and anti-bullying at elementary school: A controlled study. *Health Promotion International, 27*(1), 5–14.

*Kärnä, A., Voeten, M., Little, T. D., Alanen, E., Poskiparta, E., & Salmivalli, C. (2013). Effectiveness of the KiVa antibullying program: Grades 1–3 and 7–9. *Journal of*

Educational Psychology, 105(2), 535–551.

*Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large-scale evaluation of the KiVa antibullying program: Grades 4–6. *Child Development, 82*(1), 311–330.

Kelly, E. V., Newton, N. C., Stapinski, L. A., Slade, T., Barrett, E. L., Conrod, P. J., & Teesson, M. (2015). Suicidality, internalizing problems and externalizing problems among adolescent bullies, victims and bully-victims. *Preventive Medicine, 73*, 100–105.

Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin, 140*(4), 1073–1137.

*Krueger, L. M. (2010). *The implementation of an anti-bullying program to reduce bullying behaviors on elementary school buses* (Unpublished doctoral dissertation). D'Youville College.

Lee, S., Kim, C-J., & Kim, D. H. (2015). A meta-analysis of the effect of school-based anti-bullying programs. *Journal of Child Health Care, 19*(2), 136–153.

Maio, G., & Haddock, G. (2014). *The psychology of attitudes and attitude change*. Sage.

McCuddy, T., & Esbensen, F-A. (2017). After the bell and into the night: The link between delinquency and traditional, cyber-, and dual-bullying victimization. *Journal of Research in Crime and Delinquency, 54*(3), 409–441.

*McLaughlin, L. P. (2009). *The effect of cognitive behavioral therapy and cognitive behavioral therapy plus media on the reduction of bullying and victimization and the increase of empathy and bystander response in a bully prevention program for urban sixth-grade students* (Unpublished doctoral dissertation). The University of Toledo.

*Menesini, E., Codecasa, E., Benelli, B., & Cowie, H. (2003). Enhancing children's responsibility to take action against bullying: Evaluation of a befriending intervention in Italian middle schools. *Aggressive Behavior, 29*(1), 1–14.

Merrell, K. W., Gueldner, B. A., Ross, S. W., & Isava, D. M. (2008). How effective are school bullying intervention programs? A meta-analysis of intervention research. *School Psychology Quarterly, 23*(1), 26–42.

*Meyer, N., & Lesch, E. (2000). An analysis of the limitations of a behavioural programme for bullying boys from a subeconomic environment. *Southern African Journal of Child and Adolescent Mental Health, 12*(1), 59–69.

Modecki, K. L., Minchin, J., Harbaugh, A. G., Guerra, N. G., & Runions, K. C. (2014). Bullying prevalence across contexts: A meta-analysis measuring cyber and traditional bullying. *Journal of Adolescent Health, 55*(5), 602–611.

*Naidoo, S., Satorius, B. K., de Vries, H., & Taylor, M. (2016). Verbal bullying changes among students following an educational intervention using the integrated model for behavior change. *Journal of School Health, 86*(11), 813–822.

*Nocentini, A., & Menesini, E. (2016). KiVa Anti-Bullying program in Italy: Evidence of effectiveness in a randomized control trial. *Prevention Science, 17*(8), 1012–1023.

Olweus, D. (1994a). Bullying at school. *Promotion & Education, 60*(6), 97–130.

Olweus, D. (1994b). Bullying at school: Basic facts and effects of a school based intervention program. *Journal of Child Psychology and Psychiatry, 35*(7), 1171–1190.

Olweus, D. (2005). A useful evaluation design, and effects of the olweus bullying prevention program. *Psychology, Crime & Law, 11*(4), 389–402.

*Palladino, B. E., Nocentini, A., & Menesini, E. (2016). Evidence-based intervention against bullying and cyberbullying: Evaluation of the NoTrap! program in two independent trials. *Aggressive Behavior, 42*(2), 194–206.

*Rawana, J. S., Norwood, S. J., & Whitley, J. (2011). A mixed-method evaluation of a strength-based bullying prevention program. *Canadian Journal of School Psychology, 26*(4), 283–300.

Saarento, S., Garandeau, C. F., & Salmivalli, C. (2015). Classroom- and school-level contributions to bullying and victimization: A review. *Journal of Community & Applied Social Psychology, 25*(3), 204–218.

*Şahin, M. (2012). An investigation into the efficiency of empathy training program on preventing bullying in primary schools. *Children and Youth Services Review, 34*(7), 1325–1330.

*Salmivalli, C., Kärnä, A., & Poskiparta, E. (2011). Counteracting bullying in finland: The KiVa program and its effects on different forms of being bullied. *International Journal of Behavioral Development, 35*(5), 405–411.

*Solomontos-Kountouri, O., Gradinger, P., Yanagida, T., & Strohmeier, D. (2016). The implementation and evaluation of the ViSC program in cyprus: Challenges of cross-national dissemination and evaluation results. *European Journal of Developmental Psychology, 13*(6), 737–755.

*Stan, C., & Beldean, I. G. (2014). The development of social and emotional skills of students-ways to reduce the frequency of bullying-type events. Experimental results. *Procedia-Social and Behavioral Sciences, 114*, 735–743.

*Stevens, V., Bourdeaudhuij, I., & van Oost, P. (2000). Bullying in flemish schools: An evaluation of anti-bullying intervention in primary and secondary schools. *British Journal of Educational Psychology, 70*(2), 195–210.

*Trip, S., Bora, C., Sipos-Gug, S., Tocai, I., Gradinger, P., Yanagida, T., & Strohmeier, D. (2015). Bullying prevention in schools by targeting cognitions, emotions, and behavior: Evaluating the effectiveness of the REBE-ViSC program. *Journal of Counseling Psychology, 62*(4), 732–740.

Ttofi, M. M., & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce bullying: A systematic and meta-analytic review. *Journal of Experimental Criminology, 7*(1), 27–56.

Ttofi, M. M., Farrington, D. P., Lösel, F., Crago, R. V., & Theodorakis, N. (2016). School bullying and drug use later in life: A meta-analytic investigation. *School Psychology Quarterly, 31*(1), 8–27.

*van der Ploeg, R., Steglich, C., & Veenstra, R. (2016). The support group approach in the Dutch KiVa anti-bullying programme: Effects on victimisation, defending and well-being at school. *Educational Research, 58*(3), 221–236.

Vreeman, R. C., & Carroll, A. E. (2007). A systematic review of school-based interventions to prevent bullying. *Archives of Pediatrics & Adolescent Medicine, 161*(1), 78–88.

*Yanagida, T., Strohmeier, D., & Spiel, C. (2016). Dynamic change of aggressive behavior and victimization among adolescents: Effectiveness of the ViSC program. *Journal of Clinical Child & Adolescent Psychology*, 1–15.

Yang, A., & Salmivalli, C. (2015). Effectiveness of the KiVa antibullying programme on bully-victims, bullies and victims. *Educational Research, 57*(1), 80–90.

Yeager, D. S., Fong, C. J., Lee, H. Y., & Espelage, D. L. (2015). Declines in efficacy of anti-bullying programs among older adolescents: Theory and a three-level meta-analysis. *Journal of Applied Developmental Psychology, 37*(1), 36–51.

Zeng, X., Zhang, Y., Kwong, J. S., Zhang, C., Li, S., Sun, F., … Du, L. (2015). The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: A systematic review. *Journal of Evidence-Based Medicine, 8*(1), 2–10.

Zych, I., Ortega-Ruiz, R., & Del Rey, R. (2015). Systematic review of theoretical studies on bullying and cyberbullying: Facts, knowledge, prevention, and intervention. *Aggression and Violent Behavior, 23*, 1–21.

The effect and moderators of school-based anti-bullying programs: Meta-analysis and GRADE evidence

ZHAO Lingbo[1]; LAI Lizu[2]; LIN Yuzhong[1]; ZHAO Chunxiao[2]; REN Zhihong[2]

([1] *School of Humanities and Social Sciences, Fuzhou University, Fuzhou* 350108, *China*) ([2] *Key Laboratory of Adolescent Cyberpsychology and Behavior (CCNU), Ministry of Education; School of Psychology, Central China Normal University; Key Laboratory of Human Development and Mental Health of Hubei Province, Wuhan* 430079, *China*)

Abstract: School bullying has become a public health issue in the past several decades which caused physical and psychological harm on bullies, victims as well as bystanders. Therefore, a series of anti-bullying programs were designed to reduce the bullying or the harmfulness of bullying, but the effectiveness of these programs was inconsistent. Forty-three pre-post randomized controlled trail studies were selected to examine the effect of school-based anti-bullying programs. The evidence quality was screened under the GRADE protocol. Results suggested a significant effect for anti-bullying programs which targets bullies (g = 0.57, p < 0.05), while those who target victims, victimization decreased (g = 0.40, p < 0.05), and overall metal health improved (g = 0.40, p < 0.01). The pooled effect size indicated that school-based anti-bullying programs have a small effect on bullying and victimization (g = 0.17, p < 0.001) and influenced by publication bias. Attitude changed moderately (g = 0.19, p < 0.001). The evidence of attitude change was moderate, other outcome variables have a very low or low evidence quality according to GRADE protocol. For the programs targeting universal students, programs which was less than one semester performed better, and the global intervention showed more effective than education. The effect size of anti-bullying programs targeting bullies or victims was moderate, but more research is needed. Though school-based anti-bullying programs which target universal students have a relatively small effect size, however, they are practical for real-world applications. The dose of intervention and the feature of programs have influence on the effect of anti-bullying programs.

Key words: school anti-bullying; intervention effect; meta-analysis; GRADE